Before the
FEDERAL COMMUNICATIONS COMMISSION
Washington, D.C. 20554

In the Matter of	)	
	)	
Broadband Industry Practices	)	WC Docket No. 07-52
	)	
	)	

TESTIMONY OF GEORGE OU

FCC Broadband Industry Practices Hearing
Stanford University
April 17, 2008

George Ou
[george_ou@LANArchitect.net]

Table of contents

Opening statement

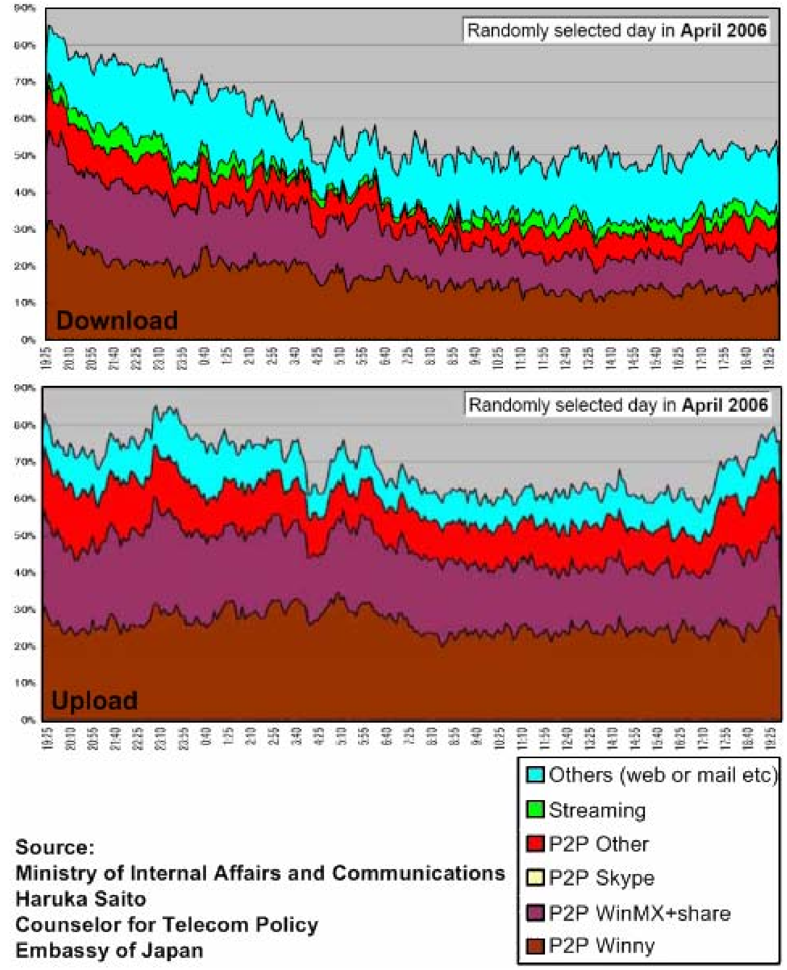
Good afternoon Mr. Chairman and Commissioners, my name is George Ou and I am a former Network Engineer who designed and built network and server infrastructure for Fortune 100 companies. Thank you for inviting me here to speak to you about this critical matter of Network Management. I'm here to explain why network management has and always will be an essential part of the Internet.

Past Internet crises

The Internet is an evolving standard that had its share of growing pains and it continues to face them today. The rise of FTP (File Transfer Protocol) caused "congestion collapse" in 1986 because the Internet at the time lacked a functional congestion avoidance mechanism. This was patched in 1987 by Van Jacobson's congestion avoidance algorithm which fairly allocated network resources to the applications of its time. Despite newer and more efficient congestion control standards, Jacobson's algorithm has remained dominant for over 20 years.

By the mid 1990s, the rise of the web browser turned the "World Wide Web" in to the "World Wide Wait". Because the first web browsers weren't optimized for resource utilization, they were tuned to be more resource friendly in the re-write of HTTP between version 1.0 and 1.1.

Today's crisis on the Internet



Today we face a whole new problem. P2P (Peer-to-Peer) applications are causing a new congestion collapse because broadcast video is migrating to VOD (Video on Demand) over the Internet. Because of loopholes in the TCP standard, a small percentage of users utilizing P2P are appropriating the vast majority of network capacity at the expense of every other consumer. Interactive applications like web browsing or real-time applications like Voice over IP are being unnecessarily degraded not by carriers, but by P2P users.

Even Japan, with one of the world's fastest Internet broadband infrastructure, where many homes have access to 100 Mbps or even 1000 Mbps fiber service are facing the same P2P induced congestion collapse. This has caused their government to spend the last two years studying the issue. Recently, the Japanese Ministry of Internal Affairs and Communications released a study showing that just 1% of Japan's broadband users using P2P account for roughly 47% of Japan's Internet usage. Furthermore, only 10% of Japan's broadband users using P2P account, on average, for 75% of all Internet usage.

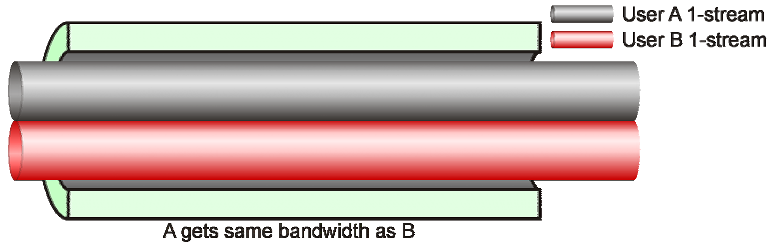
This would be like 10 passengers on a 100-seat airplane taking up 75 seats and forcing the other 90 passengers to share the remaining 25 seats. One of those 10 passengers would take up 47 seats while the other 9 take up 28 seats. While it's perfectly acceptable for 10 people to take 75 seats when there are only 25 other people on the plane who don't mind using a single seat, it is outrageous when 90 passengers who paid just as much for their ticket as the 10 seat hogs get "de-prioritized" in to 25 seats.

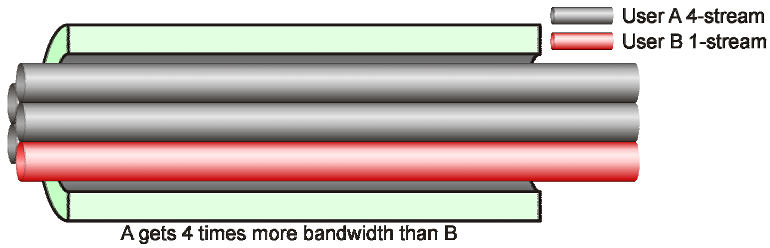
On the airplane, the obvious thing to do is to tell those people to take their feet and bags off the other seats so that other passengers may sit down. Yet on the Internet because it isn't easy for everyone to see what is really going on, the P2P bandwidth hogs yell "discrimination" and persuade activists to portray them as the victims of evil corporations who are being deprived of their civil rights. If anyone dares to throttle their over overconsumption in any way, activist groups demand trillion-dollar FCC fines and immediate enjoinments[1] before the facts are even in. But there's nothing neutral or fair about what these groups are asking for and they're not the protectors of consumer rights they portray themselves to be.

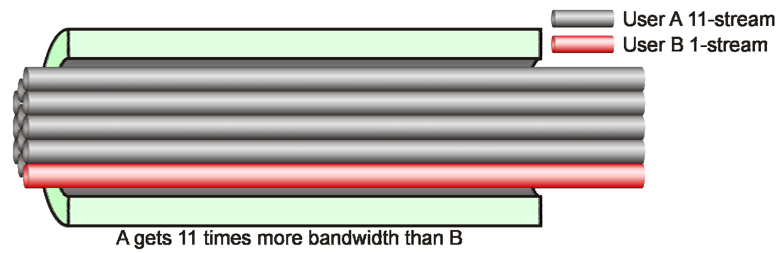
1) http://www.freepress.net/files/fp_pk_comcast_complaint.pdf

The unfairness of TCP congestion control

So how could such an extreme state of unfairness be possible when TCP congestion control was supposedly fair? It's because Jacobson's TCP algorithm was only designed for the FTP problem of the 1980s, not the problems of the 1990s and beyond. It had a major loophole because it only sought to balance the flow rate of each TCP stream with no regard to how many TCP streams a person's application could open. While Jacobson's TCP algorithm worked for early Internet applications that only used one active TCP stream at a time, it's completely ineffective for the applications of today.



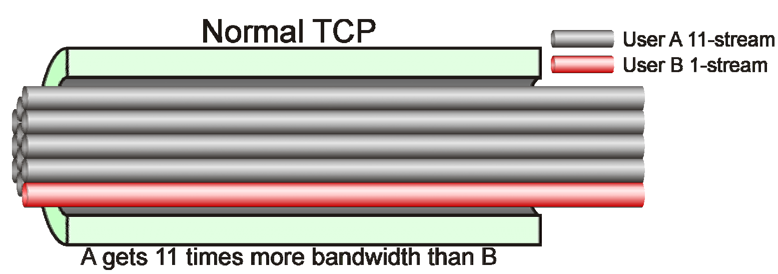


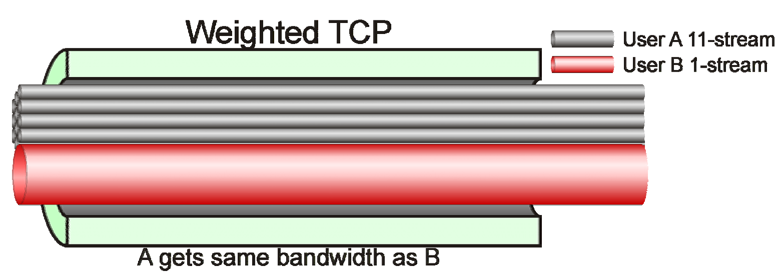


By 1999, the first P2P application called "Swarmcast" appeared on the Internet and began fully exploiting that loophole. When a P2P application opens up multiple TCP streams, each TCP stream flows at the same rate of TCP stream. That means a person using 4 TCP streams gets 4 times more bandwidth on a congested network than a person using the normal single TCP stream. The person using 11 TCP streams will get 11 times more bandwidth. Coupled with the fact that P2P applications operate non-stop throughout the day while normal applications only transmit and receive data a small percentage of the time, in short bursts, it is easy to see how only 1% of users, using P2P, have the ability to consume nearly half of all resources.

Fixing the unfairness of TCP

Thanks to the work of Frank Kelly and BT chief researcher Bob Briscoe, the matter of TCP unfairness is now before the IETF (Internet Engineering Task Force). Briscoe has issued an official IETF problem statement and he wants close the loopholes in TCP.

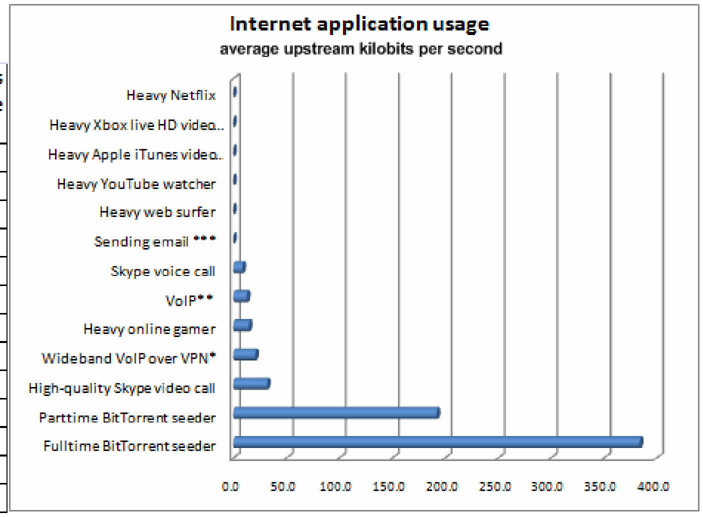




Unfortunately, making fundamental changes at the TCP standards level is extremely difficult and time consuming. Even when superior TCP congestion control mechanisms like ECN (Explicit Congestion Notification) were standardized in 2001, we have yet to see any real-world deployments because we can't force the billion-device installed base to upgrade. We're no longer in a time when a few thousand system administrators could patch all 30,000 computers on the entire Internet with Jacobson's algorithm in 1987.

The effect of non-stop application usage on base loads

	Hours used per day	In-use upstream kbps	Average upstream kbps	Kilobytes sent per day	Sessions in use
Fulltime BitTorrent seeder	24	384	384.0	4,147,200	~20
Parttime BitTorrent seeder	12	384	192.0	2,073,600	~20
High-quality Skype video call	2	384	32.0	345,600	1
Wideband VoIP over VPN*	5	98.8	20.6	222,300	1
Heavy online gamer	10	35	14.6	157,500	1
VoIP**	6	50.8	12.7	137,160	1
Skype voice call	6	34.8	8.7	93,960	1
Sending email ***	0.0031	384	0.05	535	1
Heavy web surfer	12	0.1	0.05	<500	1
Heavy YouTube watcher	12	0.1	0.05	<500	1
Heavy Apple iTunes video user	12	0.1	0.05	<500	1
Heavy Xbox live HD video user	12	0.1	0.05	<500	1
Heavy Netflix	12	0.1	0.05	<500	1



* Corporate VPN telecommuter worker using G.722 codec @ 64 kbps payload and 33.8 kbps packetization overhead
** Vonage or Lingo SIP-based VoIP service with G.726 codec @ 32 kbps payload and 18.8 kbps packetization overhead
*** I calculated that I sent 29976 kilobytes of mail over the last 56 days averaging 0.04956 kbps

In the real world, it's extremely unlikely that there would be a BitTorrent or P2P seeder who isn't a bandwidth hog. Anyone saturating the upstream nonstop and using multiple TCP streams is by definition hogging a much larger share of bandwidth than anyone else. Even if a P2P user manually sets his/her bandwidth consumption down to a fraction of the peak upload and only seeds for 12 hours a day, the amount of upstream resources used still dwarfs every other application. As I show in the chart above, just the persistence advantage alone allows a BitTorrent client to take an order of magnitude more upstream bandwidth than even the next heaviest upstream application.

Just as it is possible to "game " the TCP algorithm using P2P applications, it is also possible to construct a contrived example in which someone uses a P2P application to upload a 4.23 MB copy of the King James Version of the Holy Bible in order, they posit, to distribute the Bible to a single person. Then, when the downloading party tries to obtain the content with enough trials, they eventually trigger Comcast's network management response (which is not accustomed to seeing the anomalous situation in which a P2P application is used to exchange information between only two users). The "testing" party can then claim that even a minimal use of upstream bandwidth by a P2P application was targeted unfairly by Comcast's TCP resets. On the face of it, this claim sounds plausible and might suggest that something is wrong. But the minute you realize that Comcast gives you a gigabyte of web hosting space which allows you to distribute to many times more people at more than 20 times the speed of using a BitTorrent seed without the need to tie up your personal computer and broadband connection, you realize the implausibility of this scenario as a practical matter.

To illustrate this point, Richard Bennett used his Comcast-allocated web space to host that copy of the King James Bible at http://home.comcast.net/~richard_bennett/site/ which is far more accessible to anyone with a web browser at blazing fast speeds. This not only benefits the users because it's easy and fast, it also alleviates congestion on the scarce upstream path on a cable broadband provider's last-mile DOCSIS network.

Limitations of shared-medium broadband networks

	Upstream	**Downstream**	**# of BitTorrent 24×7 seeders to kill unmanaged network**
Cable DOCSIS 1.1	10 Mbps	40 Mbps	Less than 26 (1)
Cable DOCSIS 3.0 (4)	30 Mbps	80 Mbps	Less than 30 (2)
Wireless 802.11b ISP	4 to 6 Mbps shared		One (3)

1. Fewer than 26 fulltime P2P seeders saturating their upstream at 384 kbps 24×7 kills an unmanaged DOCSIS 1.1 network.
2. Fewer than 30 fulltime P2P seeders saturating their upstream at 1 Mbps 24×7 kills an unmanaged DOCSIS 3.0 network.
3. One fulltime P2P seeders OR P2P uploaders/downloaders can kill an unmanaged Wireless 802.11g network.
4. First implementations of Comcast's DOCSIS 3.0 use 2 bonded-channels downstream and a single upstream channel.

How do we deal with the immediate crisis?

The reality is that we cannot expect any meaningful fixes on TCP for the foreseeable future, let alone world-wide deployment of such a remedy. The solution must be found in the network itself because we cannot expect ordinary users to patch their computers to behave more politely even if a patch was approved and standardized. The network must implement and enforce fairness or else we have a state of anarchy where the "wants" of the few constrain the majority of the capacity that was intended for all paying users.

But as with any technology, there are growing pains in network-managed fairness. The first generation solutions would typically use protocol throttling techniques to neutralize the disproportionately large amounts of bandwidth that P2P applications consume. Throttling is not always the most accurate method and it can't work in all network topologies but it is cheap to deploy and used where it is effective. Other first-generation solutions like the Sandvine appliance used by cable broadband companies had occasional false positives like the accidental blockage of IBM Lotus Notes, but those bugs were quickly fixed as soon as they were identified.

The bigger problem with these first generation solutions is that they may not identify all of the bandwidth hogs and they can be fooled by protocol obfuscation techniques. It is also less than ideal to force a complete stoppage of a P2P BitTorrent seed even if it's only for a few minutes at a time during congested periods of the day rather than just slowing them down. But slowing down a specific bandwidth hog requires more drastic in-line changes to the network infrastructure, as opposed to the out-of-band Sandvine boxes that merely issued TCP resets to occasionally stop P2P seeding.

> Note: TCP resets are a commonly implemented feature in Internet firewalls and routers. TCP resets operate on the "control bits" in the transport layer and they do not constitute any modification or forgery of user data.

Critics like the Free Press and EFF claim that it would be better if users controlled their own throttling through economic incentives in the form of metered Internet access. But that would be far more draconian since users have to manually shut down P2P completely for 12 hours a day compared to having the ISP automatically shut down P2P seeds a few minutes at a time while not affecting normal P2P uploading and downloading. Any parent would appreciate the risk of their teenage child racking up thousand-dollar broadband bills because they thought it would be cool to try a new application that happens to conscript their computer as a P2P file distribution seed.

The immediate crisis is even more problematic for the wireless industry where spectrum and bandwidth is even scarcer and the capacity shared between more people than a cable DOCSIS network. One option in use today is to tell customers upfront that P2P applications aren't supported. Small wireless operators like LARIAT operating in the unlicensed 2.4 GHz spectrum space have so little shared capacity and their backhaul connections to the Internet are so expensive that they can't afford to have uploading bandwidth costs of other content corporations shifted on to them. Not only are the bandwidth costs shifted to them under the P2P distribution model, the costs are amplified by an order of magnitude because bandwidth out in the rural areas cost far more than the concentrated bandwidth in the data centers. These small wireless ISPs are often times the only Internet Service Providers in their space because larger corporations don't want to serve these less lucrative areas. As Dr. Robert D. Atkinson and Philip J. Weiser of ITIF argued[2], "Even port blocking, for example, might be defensible under certain circumstances." In this case, having a small ISP that blocks P2P serve an area is better than none at all or better than one that charges by the bit.

2) A "Third Way" on Network Neutrality http://www.itif.org/files/netneutrality.pdf

Next generation network management technologies

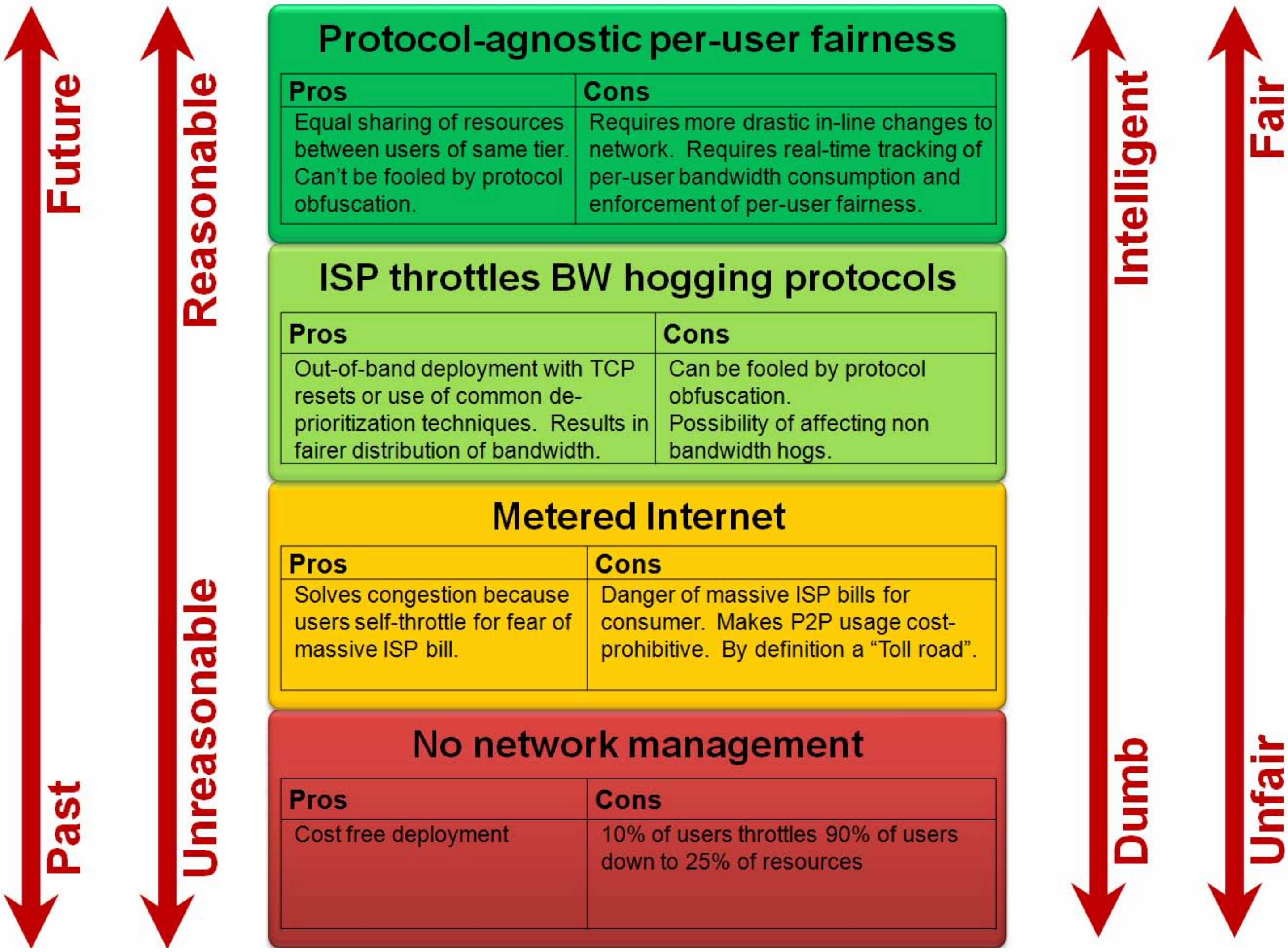
As technology matures and network management becomes more sophisticated, companies that are most susceptible to P2P congestion (those using cable or wireless shared media transmission) are committing to newer in-line technologies that are protocol-agnostic and enforce real-time, per-user bandwidth fairness. These newer technologies will avoid false positives or false negatives and more importantly, they can't be fooled by protocol obfuscation techniques because they only look at bandwidth usage patterns. They also respond to bandwidth hogs by slowing them down rather than occasionally stopping them with TCP resets.

Comcast's CTO Tony Werner explained that any congested link on their network will have 2% of the users taking approximately half of all capacity at the expense of everyone else and this is consistent with the data from Japan's government. The new in-line technology Comcast is experimenting with now will de-prioritize and throttle that 2% of bandwidth hogs down to roughly 25% of all resources and that will be enough to alleviate congestion for everyone else. When the network isn't congested, those few P2P users can resume taking as much as they please.

BitTorrent Corporation knows that the metered Internet service plans promoted by Free Press will severely harm use of their application and their business model, which is why they know it's in their best interest to work with ISPs. BitTorrent will begin to clearly label their own traffic as P2P instead of obscuring their protocol headers like most P2P clients so that network operators can properly manage them. BitTorrent is also working closely with the P4P working group to make P2P clients more efficient by selecting nearby peers instead of random peers to minimize load on the network. Finally, BitTorrent K.K., BitTorrent's Japanese subsidiary, was part of a recent working group of ISPs and applications developers in Japan that has developed "best practices" guidelines for traffic shaping—recognizing that the mere addition of network capacity will not correct a problem created by an application designed to consume all available capacity.

How do we define what is "reasonable network management"?

It should be reasonable to conclude that "fairness" to each paying consumer should be a key criterion in determining how reasonable a solution is. Based on that criterion, I worked out the chart below with the solutions that are either in use or being proposed.



I've defined **"No network management"** as the <u>most</u> unreasonable form of network management because it allows 10% of the users to throttle 90% of the users by leaving them with 25% of the resources. What does it matter that it's a small minority of users doing the "throttling" here rather than some corporation? Isn't the greatest amount of harm being perpetrated on the largest number of paying consumers under a non-managed network?

The Free Press argues that protocol-agnostic per-user fairness, Metered Internet, or No network management (random packet drops) would all be superior to the protocol throttling methods that employ de-prioritization or TCP reset and that protocol throttling should be banned. While we can all agree that we should move to the protocol-agnostic per-user fairness model of network management, it would be harmful to the vast majority of consumers if we insisted on immediate perfection today and banned something that was mostly fair and mostly reasonable. The market is already moving to the newer more advanced network management techniques and we should give that a chance to mature rather than forcing a change for the worst in the meantime.

Network management ensures a harmonious coexistence

As Japan's broadband experience has shown, we will never grow our way out of congestion and we will never have enough bandwidth. But managing a network can ensure a harmonious coexistence where P2P background applications and traditional interactive or real-time applications all get what they want. All too often, non-technical Network Neutrality proponents confuse network priority with network throughput when they in fact have nothing to do with each other.

P2P will always consume the lion's share in traffic volume but they can have all the volume in the world even when they have the least priority. Interactive and real-time applications fundamentally have low volume requirements and there's no reason they shouldn't get maximum priority. It makes no difference to a P2P application if you send the interactive or real-time packets in a quick priority burst since that would only get them out of the way sooner. The interactive or real-time applications only have a small and fixed amount of data to transport so the amount they can displace the larger P2P file transfers remains constant regardless of priority given to the smaller payloads. The P2P file transfer will still be completed in the same amount of time whether that's an hour or a day. But receiving priority makes all the difference in the world to the interactive and real-time application because a human is waiting for an immediate response and voice or gaming applications can't tolerate delays that are more than a tenth of a second.

So with network management, background applications like P2P can get all the traffic volume they want in the same timely fashion with no need for draconian metered pricing. Interactive or real-time applications get the priority they need so they don't get drowned out P2P applications. But without network management regardless of how much capacity you throw at the problem, P2P applications won't run much faster but they will drown out the interactive and real-time applications on the Internet.

Conclusion

The purpose of my testimony is for the Commission to recognize the efforts the networking industry is making to improve network traffic management and consumer disclosure. The industry faces a hard up-hill climb to educate consumers about the services they offer, their limitations, and their terms of use because we aren't a nation of network engineers. Perhaps the worst side-effect of the current controversy over P2P management is the large number of false and misleading statements about network engineering that have become part of the public debate. The Commission should be particularly mindful of its role in educating the public toward a genuine understanding of the Internet and its unique properties compared to the traditional telephone network. Networking is a difficult and complicated subject and I hope the Commission will consider the pragmatic, engineering concerns when crafting its policy solutions. Chairman Martin, and the other members of the Commission, thank you for taking the time to listen.

Sincerely,
George Ou
Former Network Engineer

Acknowledgements
I'd like to thank Network Architect Richard Bennett, who testified at the FCC hearings at Harvard, for contributing some of the ideas and edits to this paper. I'd also like to thank BT chief researcher Bob Briscoe for his influence on this paper as well and anyone else who helped edit this paper.